CS 134: Networks	Lecture 7 **outline**, Monday, 2/13
Prof. Yaron Singer	Spring 2017

1 Overview

Question: *how do natural sample biases shape our perception of the world?*

1.1 Example: Everyone at the Gym Looks Fitter than Me

The gym paradox: *The average fitness of the people you see in the gym is higher than the average fitness of the gym's members.*

Consider gym with 5 members $\{0, 1, 2, 3, 4\}$; member i spends i hours distributed uniformly at random. Suppose we walk into the gym at a random hour.

- What is the average fitness of gym members?
- What is the average fitness of someone at the gym when I visit?

1.2 Example: I am always in classes bigger than average

The class size paradox: *Students experience an average class size larger than the school's average class size.*

Consider a department with two classes; every student in the department takes one or the other, but not both. One is a large introductory course with 90 freshmen in it. The other is an advanced seminar with 10 seniors.

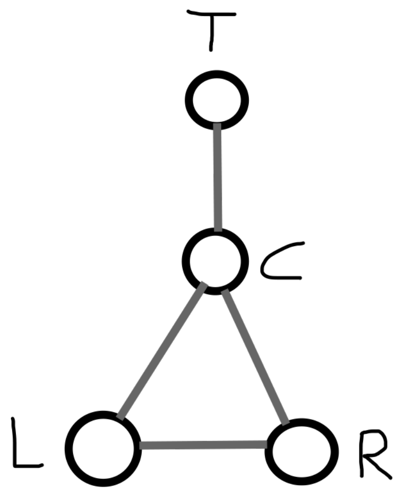
- What is the average class size?
- What is the class size that a student experiences on average?.

2 The Friendship Paradox: Why Your Friends Have More Friends than You

Theorem (The Friendship Paradox)**.** *In any graph, the average degree of nodes is no more than the average degree of all the neighbors.*

The term "the friendship paradox" and the above proof are due to Scott Feld [1]. Ideally, to make a statement like "your friends have more friends than you", we would like to prove something a bit stronger, namely that for every node its degree is bounded from above by the average degree of its neighbors.

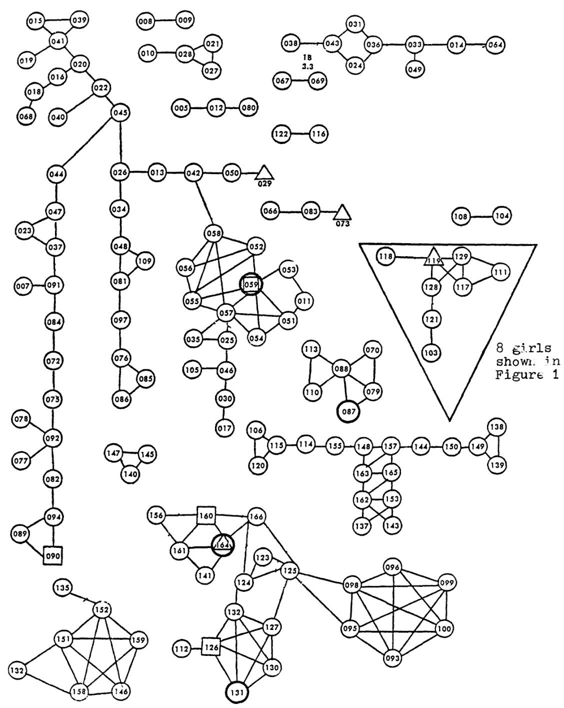
Small example. Consider the following network:

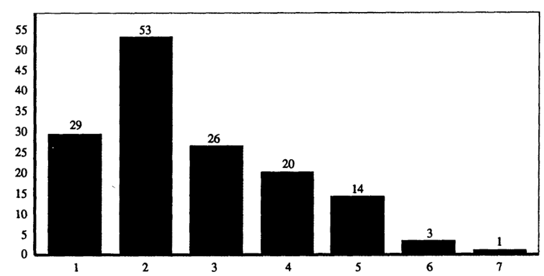


- The average degree of individuals is $\frac{1+3+2+2}{4} = 2$.
- The average degree of friends is $\frac{1+3*3+2*2+2*2}{8} = \frac{9}{4}$. To see this, note that
  - The friend of T has degree 3.
  - The three friends of C have degrees 1, 2 and 2.
  - The two friends of L have degrees 2 and 3.
  - The two friends of R have degrees 2 and 3.

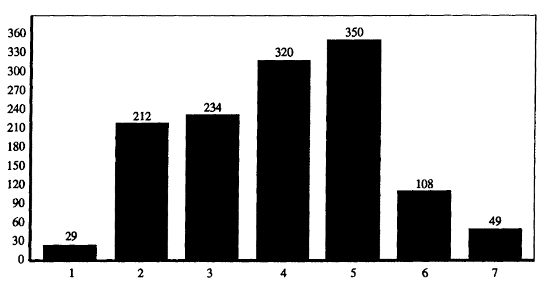
Larger example. In The Adolescent Society, Coleman (1961) collected data on friendships among the students in 12 high schools. Individuals were asked to name their friends, and pairs of individuals who named one another were given particular attention. It is these "friendships" that will be used here.

To illustrate the phenomenon under study here, consider the set of relationships depicted in Figure 2, the complete network of all of the girls in "Marketville", one of the High Schools in the study. Figure 3(a) depicts the distributions of friends of individuals and Figure 3(b) the distribution of friends of friends. Friends tend to have more friends than individuals.





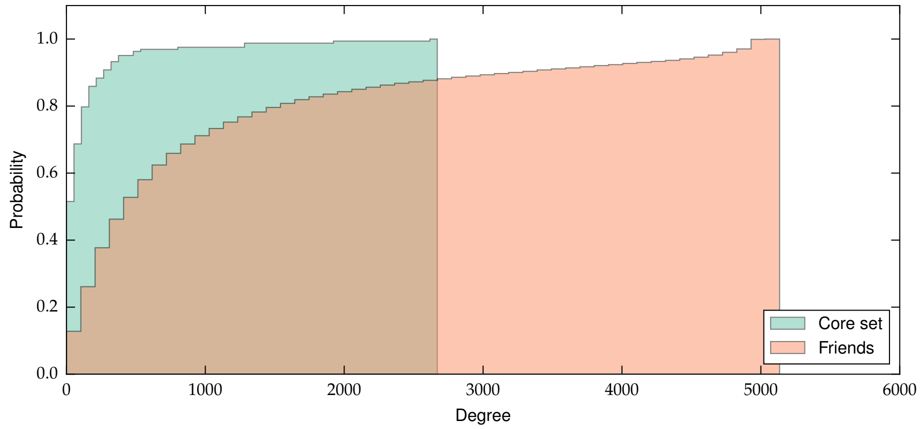
a) The mean is 2.7.

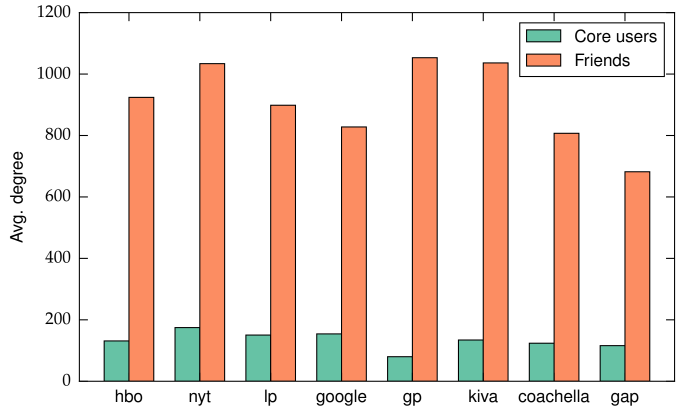


b) The mean is 3.4

(*a*) Distribution of numbers of friends for Marketville girls; (*b*) distribution of number of friends' friends for Marketville girls.

2.1 Friendship paradox in Facebook Pages





3 Survivor Bias: Abraham Wald and the Missing Bullet Holes

Section of Plane	Bullet holes per square foot
Engine	1.11
Fuselage	1.73
Fuel system	1.55
Rest of plane	1.8

Section of Plane Hit	Number of Planes Hit	Number of Planes Back
Engine	20	2
Fuselage	20	9
Fuel system	20	6
Rest of plane	20	10

References

[1] Feld, Scott L. *Why your friends have more friends than you do*, American Journal of Sociology, pages 1464–1477, 1991,JSTOR